REMARKS

[0003]   Applicant respectfully requests reconsideration and allowance of all of the claims of the application. Claims 1-7, 15-19, 64, 65, and 67-75are presently pending.   Claims amended herein are 1-4, 15-17, 64, 65, and 67. Claims withdrawn or cancelled herein are none.   New claims added herein are 72-75.

Statement of Substance of Interview

[0004]   The Examiner graciously talked with me—the undersigned representative for the Applicant—on May 2, 2007. Applicant greatly appreciates the Examiner's willingness to talk. Such willingness is invaluable to both of us in our common goal of an expedited prosecution of this patent application.

[0005]   During the interview, I discussed how the claims differed from the cited art, namely Hull. Without conceding the propriety of the rejections and in the interest of expediting prosecution, I also proposed several possible clarifying amendments.

[0006]   Based upon the Examiner's suggestions, I amend herein several claims to clarify that the filtered subtext being hashed is in an "image." In other words, rather than hashing the text or some "descriptor" of that text, *an image of the text* is being hashed. This differs from Hull. I will elaborate more below in the discussion of the § 103 rejections.

[0007] Applicant herein amends the claims in the manner in accordance with what was discussed during the interview. Accordingly, Applicant submits that the pending claims are allowable over the cited art of record for at least the reasons discussed during the interview.

Formal Request for an Interview

[0008] If the Examiner's reply to this communication is anything other than allowance of all pending claims, then I formally request an interview with the Examiner. I encourage the Examiner to call me—the undersigned representative for the Applicant—so that we can talk about this matter so as to resolve any outstanding issues quickly and efficiently over the phone.

[0009] Please contact me or my assistant to schedule a date and time for a telephone interview that is most convenient for both of us. While email works great for us, I welcome your call to either of us as well. Our contact information may be found on the last page of this response.

Claim Amendments and Additions

[0010] Without conceding the propriety of the rejections herein and in the interest of expediting prosecution, Applicant amends claims 1-4, 15-17, 64, 65, and 67 herein.

[0011] Furthermore, Applicant adds new claims 72-75 herein. These new claims are fully supported by Application and therefore do not constitute new matter. For example, support for removal of superfluous content and canonical format (e.g., claim 72) can be found in the Application on p. 11, line 3 through

p. 12, line 12 and support for pseudo-random extraction (e.g., claim 73) is found at p. 12, lines 14-21.

Substantive Matters

Claim Rejections under §103

[0012]   The Office rejects all of the pending claims under §103. For the reasons set forth below, the Office has not shown that the cited references disclose, teach or suggest (under §103) the rejected claims. Accordingly, Applicant respectfully requests that the rejections be withdrawn and the case be passed along to issuance.

[0013]   The Office's rejections are based upon Hull [Hull et al., U.S. Patent No. 5,465,353 (issued 11/7/1995)] and Pugh [Pugh et a., U.S. Patent No. 6,658,423].

Overview of the Application

[0014]   The Application describes a technology for recognizing the content of text documents. The technology may detect similarity between text-based works in an automatic and accurate manner. Furthermore, it may categorize content of text-based works in an automatic and accurate manner.

[0015]   Generally, the technology determines one or more hash values for images of the content of a text document. Furthermore, the technology may generate a "sifted text" version of a document.

[0016] In one implementation described herein, document recognition is used to determine whether the content of one document is copied (i.e., plagiarized) from another document. This is done by comparing hash values of documents (or alternatively their sifted text).

Hull

[0017] The Office cites Hull as its primary reference in its obviousness-based rejections. Hull discloses a document matching and retrieval system where an input document is matched against a database of documents, using a descriptor database which lists descriptors and points to a list of documents containing features from which the descriptor is derived document. The descriptors are selected to be invariant to distortions caused by digitizing the documents or differences between the input document and its match in the document database. An array of accumulators is used to accumulate votes for each document in the document database as the descriptor base is scanned, wherein a vote is added to an accumulator for a document if the document is on the list as having a descriptor which is also found in the input document. The document which accumulates the most votes is returned as the matching document, or the documents with more than a threshold number of votes are returned.

Serial No.: 09/843,255
Atty Docket No.: MS1-0647US
Atty/Agent: Kasey C. Christie
RESPONSE TO NON-FINAL OFFICE ACTION

15

lee&hayes   The Business of IP ™
www.leehayes.com   509.324.9256

Obviousness Rejections Based upon Hull and Pugh

[0018] The Office rejects all pending claims (claims 1-7, 15-19, 64-65 and 67-71) under USC § 103(a) as being unpatentable over Hull in view of Pugh. Applicant respectfully traverses the rejections of these claims.

[0019] In light of the claim amendments made herein, Applicant submits that all of the outstanding substantive claim rejections are moot.

[0020] In particular, no cited reference discloses (with emphasis added), "deriving a hash value representative of the filtered subtext, perceptually similar filtered subtexts having proximally similar hash values, wherein the deriving comprises **hashing the image-based formatted, filtered subtext resulting from the formatting**" (from Claim 1).

[0021] Unlike Hull, amended claim 1 (for example) recites hashing an image—in particular, the "image-based formatted, filtered subtext." Instead of hashing an image, Hull discloses hashing "descriptors" of the content of a scanned document (such content could include text). Hull's descriptors are not an image-based format of textual content.

[0022] The following exemplary citations from Hull supports my assertion that Hull hashes descriptors and not an image:

- "...the descriptors are...hashed..." Col. 6, line 19;

- "...coupled to descriptor database 116 to provide hashed descriptors for the document to descriptor database 116..." Col. 6, lines 48-49;

- "Once the descriptors are extracted from the digital representation of the document, they are processed by hash generator 114..." Col. 7, lines 5-7;

- "...hashing said descriptors..." Claim 13;

- The description of the "Example Implementation" starting in Col 13 explains that the hash generator produced hashed descriptors.

[0023] The following exemplary citations from Hull explains what Hull's descriptors are:

> "Feature extractor 110 produces descriptor elements from the digital representation. The descriptors are preferably invariant of translation, rotation, scaling, format, font and subdivision, so that a given descriptor which is extracted from a document would still be extracted, even if the document is scanned in a different orientation, is resized or otherwise distorted. For example, with a text document, a descriptor would not be dependent on the particular margin settings or fonts used in the document. An example of specific descriptors is given below." Col. 6, line 62 through col. 7, line 4.

From the above cited text, the descriptors cannot be an "image" of textual content because an "image" textual content does not possess the characteristics of a descriptor provided by Hull. Furthermore, the description of the "Example Implementation" (starting in Col 13) discusses uses of actual descriptors. From these citations, it can be seen the Hull's descriptors are not images of textual content.

[0024] While hashing itself is not new and hashing of images is not new (see Hull, col. 2, line 47 through col. 4, line 22), hashing of an "image-based formatted, filtered subtext" of the textual content of a body of text is new and not found in any cited references.

[0025]   Consequently, Hull does not disclose all of the claimed elements and features of these claims and in particular, does not disclose "hashing the image-based formatted, filtered subtext." Accordingly, Applicant asks the Office to withdraw its rejection of these claims.

Dependent Claims

[0026] In addition to its own merits, each dependent claim is allowable for the same reasons that its base claim is allowable. Applicant requests that the Examiner withdraw the rejection of each dependent claim where its base claim is allowable.

Conclusion

[0027] All pending claims are in condition for allowance. Applicant respectfully requests reconsideration and prompt issuance of the application. If any issues remain that prevent issuance of this application, the **Examiner Vik is urged to contact me before issuing a subsequent Action**. Please call/email me or my assistant at your convenience.

Respectfully Submitted,

Dated: 5-17-07

By:
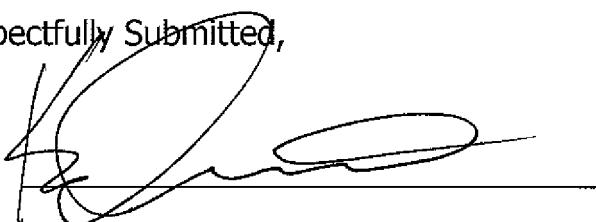
Kasey C. Christie
Reg. No. 40559
(509) 324-9256 x232
kasey@leehayes.com
www.leehayes.com

My Assistant: Carly Taylor
(509) 324-9256 x264
carly@leehayes.com

Serial No.: 09/843,255
Atty Docket No.: MS1-0647US
Atty/Agent: Kasey C. Christie
RESPONSE TO NON-FINAL OFFICE ACTION

19

lee&hayes   The Business of IP ™
www.leehayes.com   509.324.9256